DOCUMENT RESUME

ED 294 900 TM 011 552

AUTHOR Zuman, John P.
TITLE The Effectiveness of Special Preparation for the SAT: An Evaluation of a Commercial Coaching School.
PUB DATE Apr 88
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Black Students; Comparative Analysis; *Grade 11; High Schools; Hispanic Americans; Mathematics Tests; Minority Groups; *School Effectiveness; *Test Coaching; Test Results; Tutorial Programs; *Tutoring; Verbal Tests
IDENTIFIERS *Scholastic Aptitude Test

ABSTRACT

The effectiveness of a commercial preparation class in New York City for the Scholastic Aptitude Test (SAT) was studied. A group of the regular clientele of the school (55 eleventh graders) was compared with a group of low-income minority students given scholarships to the course (33 Black or Hispanic eleventh graders). The effects of the coaching school were studied, using equivalent control and experimental groups. Administrative problems and time constraints resulted in a lower quality instructional environment for the scholarship students. After coaching, regular students increased their SAT scores by about 52 points for verbal and 58 points for math. For scholarship students, coaching had no effect on verbal scores and increased their scores by about 57 points for math. This may have reflected a lack of experience in teaching minority students by the instructors as well as inadequate high school preparation for the type of instruction offered. (SLD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED294900

# The Effectiveness of Special Preparation for the SAT: An Evaluation of a Commercial Coaching School

Paper Presented at the Annual Meeting of the American Educational Research Association in New Orleans, April 1988.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JOHN P. ZUMAN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Dr. John P. Zuman

Technical Education Research Centers
1696 Massachusetts Avenue
Cambridge MA 02138

TM 011 552

## Introduction

Is coaching, or short-term preparation, effective in increasing scores on standardized tests of academic ability such as the Scholastic Aptitude Test (SAT)? This question has been intensely debated by researchers for at least the last ten years, but no clear consensus has yet emerged from the research literature. For example, Pike (1978) estimated that 21 hours of coaching on the two main math item types on the SAT would yield an average effect of 33 points on the Math portion of the SAT, and suggested that "active study directed to specific skills" (p. 61) would also significantly increase Verbal SAT scores, based on his research (Evans and Pike, 1973) and on review of the coaching literature. Alderman and Powers (1979) found an average coaching effect of 8 points on the SAT-V averaged over eight secondary schools, and concluded that coaching programs "exert little influence on actual test performance" (p. 20). Powers (1982) estimated that each of the methods for SAT preparation he studied, including coaching, had effects "of less than ten points on the 200-800 SAT score scale" (p. 1). And Slack and Porter (1980) wrote that "there is ample evidence that students can successfully train for the SAT and that the more time students devote to training, the higher their scores will be" (p. 164).

The question of the effectiveness of coaching for tests such as the SAT is a matter of concern to many besides researchers. Though published evidence on the effectiveness of coaching is ambiguous, organized efforts to coach students for the SAT and for similar tests are proliferating rapidly. Numerous high schools have incorporated SAT preparation courses into their curriculums and more and more students are enrolling in commercial test preparation centers outside of school. Some test preparation centers guarantee SAT score increases of 100 points or more for their coached students.

The coachability of the SAT has important implications for its use in the college admissions process. If students can in fact increase their scores without the requisite increases in scholastic aptitude, then the SAT can be a misleading and inappropriate indicator of potential college performance.

Moreover the question of the efficacy of coaching is directly relevant to widespread concern over the subpar average performance of Blacks, Hispanics and other minorities on standardized tests of academic skills. If coaching can be effective in significantly raising test scores of such groups, this could have important implications for future efforts to enhance minority access to post-secondary institutions that use the SAT as a criterion for admission.

The present study represents a new effort to address the effectiveness of coaching for the SAT, with particular emphasis on the effectiveness of coaching for economically disadvantaged

and ethnic minority students. Three features of the research make the results of the study salient:

1. an equivalent control group design was employed, making the study more sound than other previous studies of test coaching from the point of view of research design;

2. the study focused on a commercial test coaching program, in contrast with most previous studies, which focused on coaching and preparation programs prepared on an ad-hoc basis; and

3. the study included a sample of low-income minority students as well as a sample of the regular clientele of the coaching school.

One cause of controversy over previous studies of test coaching is that most of them employed research designs in which students were not randomly or otherwise equivalently assigned to treatment and control groups. Estimates of the effects of coaching from such studies cannot be clearly attributed to the coaching intervention, because other factors, such as self-selection and differential motivation, could account for increases in scores.

In 1978, for example, the Federal Trade Commission carried out a study of the effectiveness of two commercial test coaching programs in New York City (FTC, 1978, 1979). The performance of students who had attended one of these coaching programs was compared with that of students who apparently had not had such coaching. The students for the study were selected on a post-hoc basis from records provided by the Educational Testing Service (ETS) and the College Board. A large sample of SAT takers was selected from the New York metropolitan area and separated into a coached and non-coached group. Predictably, there were differences in background factors between the two groups, which included higher PSAT scores, higher school grades, higher family incomes and a higher percentage of private school students for the coached group.

A number of analysts have studied the data from the FTC inquiry, and have applied a variety of statistical techniques in order to control for background differences in students who did and did not attend the coaching programs, thereby attempting to isolate the effects of coaching. However Messick (1980) points out that the 20- to 30- point Verbal and Math effects for students attending one of the coaching schools is actually an estimate of the joint effects of coaching and self-selection and that it is impossible to determine with confidence whether the effects may be attributable in whole or in part to self-selection not controlled under the study design rather than to any impact of the coaching program as such.

In their meta-analytic study of previous investigations of the effects of test coaching on the SAT, DerSimonian and Laird (1983) found "a large variation in the effect of coaching from study to study that cannot be explained by sampling error" (pp. 11-13). These investigators found that matched or randomized evaluations showed a much greater degree of consistency in results (with generally smaller effects estimates) than did less well-controlled studies. One clear implication of this finding is that the larger effects of coaching reported in the literature may be indirect effects of lack of adequate controls. This hypothesis clearly speaks to the importance of using strong experimental designs in any future studies of test coaching, such as the equivalent control group design used in the present study.

However, DerSimonian and Laird also observed that the above hypothesis is tentative, because among previous studies there has been "a degree of confounding between type of evaluation and type of coaching program" (p. 14). Thus it may be that the larger effects estimates reported in the literature in the past are not merely methodological artifacts, but instead are real reflections of larger effects of better-organized, more intensive coaching programs.

Most previous studies of test coaching have focused on school-based or ad-hoc coaching programs of relatively brief duration and intensity. Indeed, only the Federal Trade Commission investigation concerning coaching for the SAT focused on a commercial coaching school. As mentioned earier, the FTC study found significant gains in Verbal and Math SAT scores for a group of students in one coaching school, but suffered from the lack of an equivalent control group.

The present study, by using an equivalent control group design and focusing on a well-organized, intensive coaching program, brings a critical piece of evidence to bear on competing hypotheses for explaining the conflicting findings of previous research on coaching. Are large effects of coaching merely artifacts of poorly controlled investigations, or are they the real results of a better-organized and more intensive coaching program?

One of the problems in interpreting the findings of the FTC study is that students attending commercial test coaching schools tend to be more affluent than the general population of students who take the SAT. This is an indirect reflection of the fact that attending commercial test coaching programs costs money - as much as $600. From the point of view of experimental design and inference, this means that using a control group of the general student population taking the SAT as controls for the self-selected group of students taking commercial coaching schools, as was attempted in the FTC study, can never completely disentangle effects of coaching from the effects of self-selection. However from a substantive point of view, the

existence of commercial test coaching schools raises some important questions with regard to educational policy and equality. If at least some commercial test coaching schools are effective in increasing students' SAT scores, and if, as the FTC study indicated, it is the more advantaged, affluent students who tend to have access to these schools, then students who are economically disadvanaged, including Black and Hispanic students, are suffering from double jeopardy in the race for college admissions. Already, from a great deal of both current and historical evidence concerning social patterns of test performance we know that economically disadvantaged and minority students tend generally to perform more poorly on tests of scholastic ability than their more advantaged peers. But if in addition they are suffering from the easier access that more affluent students have to effective coaching, then this would have a variety of possible implications for college admissions policies and efforts to promote greater equality of educational opportunity.

For these reasons, the present study examines the effects of test coaching not just for the normal population of students who can pay for the services of commercial test preparation program, but also the effects for economically disadvantaged and minority students. This was accomplished by providing scholarships for a group of such students from the Manhattan and Brooklyn public schools to attend the commercial test coaching school under study. This allows some indication of whether the effects of test coaching are any greater or less for a population of students who generally perform less well on scholastic ability tests than the effects for more advantaged students.

## The Sample

The study was conducted in the Spring of 1986 in New York City. Two groups of eleventh-grade students were sampled: one sample consisted of the regular clientele of the coaching school selected from students registered to take the coaching course (n=55), and the other sample was a group of low-income minority students who received scholarships to attend the coaching course (n=48). The regular students were largely White, from relatively wealthy homes, and many attended private schools. The scholarship students, who were either Black or Hispanic, were selected from public schools in Manhattan and Brooklyn. The Brooklyn students (n=36) were recruited by school guidance counselors from five different high schools, and the Manhattan students (n=12) were part of the "I Have a Dream" project, a program to encourage high-risk students to stay in school. Ten of the Brooklyn students and five of the Manhattan students did not complete the coaching course and are not included in the analyses.

Study Design

The two parts of the study, one involving the regular students and the other involving the scholarship students, were carried out at the same time. There were some differences between the two parts of the study, however, so they will be discussed separately.

The effect of coaching on SAT scores of regular students was studied using a quasi-experimental design. Students for this part of the study were selected in December 1985 from all students who had signed up to take the coaching course beginning either in January or March 1986. The students in the January course (n=21) were the treatment group, and students who signed up for the coaching course beginning in March (n=34) served as controls. The two groups were comparable along 19 background variables including PSAT scores, high school grades, level of motivation and self-selection into the coaching school. The PSAT taken by all regular students in the Fall of 1985 was used as the pretest. A special version of the SAT was administered to all students at the conclusion of the January coaching course, before the control group began to receive coaching. All treatment students also took the regularly administered March SAT. The control students had to take the posttest two weeks before the treatment group in order to give them time to complete the coaching course before the May SAT. However there was no contact between treatment and control students so contamination of the test is unlikely. All students received feedback questionnaires at the end of the coaching course which asked them to evaluate their experience and to mail back copies of their SAT reports. Figure 1 illustrates the overall design of the study.

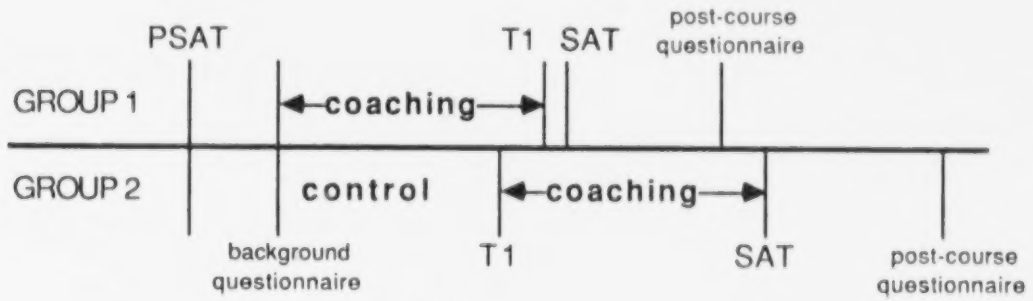



Figure 1: Overall Study Design

The effect of special preparation on the SAT scores of scholarship students was studied using an experimental design. The 48 scholarship students were randomly assigned to treatment and control groups. All scholarship students were pretested with the PSAT. At the end of the treatment all scholarship students took an unreleased version of the SAT provided by the Educational Testing Service. Nearly all students in the treatment group also took the regularly scheduled May SAT the following weekend. The control students then went on to take the coaching course and took the June SAT. The overall organization of the scholarship part of the study was very similar to the regular part of the study as shown in Figure 1. A total of 15 scholarship students did not complete the coaching course and did not take the posttests. Eight students dropped out of the treatment group and seven dropped out of the control group, but the two groups remained equivalent on all 19 available background variables despite the attrition.

## The Treatment

The coaching course for regular students consisted of three-hour classes on nine consecutive Sundays. Each class was actually two separate sessions, one for English and the other for Math, that in most cases were taught by two different instructors. Students were assigned to English and Math sessions on the basis of their pretest scores, so that students with similar scores were in the same classes. There were between 4 and 9 students in each session, with an average of 7 per class. Short breaks were scheduled between the two sessions, at which time students enjoyed snacks provided by the coaching school and interacted socially. Each Sunday class began with a group meeting at which all students listened to a short lecture by one of the staff members on SAT preparation techniques, after which the students went to their assigned sessions.

Both the Math and the English sessions stressed practice on SAT-type items using a workbook prepared by the coaching school and the ETS publication "5 SAT's". Teachers also discussed particular techniques and strategies for individual types of SAT items. The English sessions also stressed vocabulary development from a compiled list of words that were deemed likely to appear on the actual SAT. The Math sessions included discussions of logical reasoning, estimation and use of figures. Classes were organized so that students would focus on one particular item type during a particular session. Students were assigned homework of solving practice items from the workbook.

In addition to the Sunday classes, students took three practice tests during the course. These tests were taken on Saturday mornings and were intended to simulate the real SAT-taking situation. Classes following these practice test administrations were devoted to the review of the items that appeared on the tests. The tests were machine-scored after each administration and students were given detailed analyses of

their responses in order to track their progress and to learn from their mistakes. Summary reports for each class were provided for teachers to give them an accurate reflection of how their students were answering individual items.

Students also had the option of signing up for "Extra Helps" during the week. These sessions were scheduled with each student's regular teachers either in a one-to-one or a small group format. Students were eligible for these sessions if they attended the regular Sunday classes and did their homework, but still felt that they needed more help in a particular area of the test. Students who wanted still more help could sign up for individual tutoring.

The scholarship students were coached by the same instructors who taught the regular students, but in separate classes. The director of the coaching school felt that the scholarship students should be coached separately, because they scored substantially lower on the PSAT than the regular students and would therefore need a different level of preparation. The instructors had no prior experience coaching low-scoring students.

The experience of the scholarship students was somewhat different than the coaching received by the regular clients of the coaching school. First, scholarship students received only 8 weeks of instruction rather than the 9 weeks scheduled for regular students. Second, since the scholarship classes were held on Saturdays, in order not to conflict with the regular Sunday classes, the scholarship students were not given practice tests in a simulated SAT-type environment. Instead, the scholarship students took two of their practice tests during the week, administered by their school guidance counselors. The third practice test was taken by students during their regular Saturday classtime, further reducing the amount of instructional time. Third, since the scholarship students were coming from Brooklyn and Upper Manhattan to a central Manhattan location, many of them often arrived late to the Saturday classes. This was a problem for two reasons. The late-arriving students disrupted others already working, and the late students further reduced the amount of class time they received. Finally, the amount of extra help and tutoring available was severely limited due to location and limits on the teachers' time. Some of the instructors made the effort to schedule extra helps in Brooklyn and in Manhattan, and the students who attended reported these sessions as very beneficial. However most of the scholarship students did not attend these sessions, and in some cases students who said that they would come to a special help session did not show up. The net result was that teachers reduced their effort in providing extra help, particularly during the second part of the study.

The structure of the Saturday sessions was similar to the regular students' Sunday sessions. The day began at 10:30am

with a group meeting at which administrative details and test-taking techniques were discussed. After this meeting students went to their respective Math or English classes and focused their work on the coaching school workbooks and review of tests the students had taken. The scholarship classes were smaller than the regular classes - most of them had no more than four students. Between sessions the scholarship students had to buy their own lunch, as the coaching school did not provide the snacks available to their regular students. This also led to some shortening of class time, as students who left the building took their time coming back for the afternoon session. There was also a problem with the location of the classes which changed from week to week. The students had to keep track of which building their classes were in on a particular Saturday, although signs were posted to guide students to the correct location.

In summary, the administration of the coaching course for the scholarship students was different in several ways than for the regular students. Most of the differences tended to result in less instructional time for the scholarship students and a lower-quality environment for learning when compared with the coaching course provided for the regular students.

## Results

Regression analyses controlling for available background variables were employed to estimate the coaching effect. Even though there were no significant differences between treatment and control groups on any variables, it was possible that subtle differences between groups or interaction effects existed that were not detected by initial bivariate tests. The results for the regular students and the scholarship students are reported separately.

### Effects of Coaching for Regular Students

Before performing the regression analysis for regular students, the control group special SAT posttest scores had to be adjusted in order to be equivalent to the March SAT taken by the coached students. The 21 coached students took both the special SAT and the official SAT on consecutive weekends. The correlation between the two tests was very high (.87 Verbal, .81 Math), but students' mean scores on the special SAT were higher than their official SAT means by 6 Verbal points and 20 Math points. In order to compensate for the difference in means I subtracted 6 Verbal points and 20 Math points from the special SAT scores of the control students. This adjustment was necessary because it was not feasible for the control students to take the March SAT before they were coached.

An analysis of covariance model to estimate the effect of coaching on Verbal and Math SAT scores was specified. I entered

the background variables into the model to determine if any of them influenced the post-test SAT once the PSAT score and group membership were accounted for. No background variables were significant in affecting the prediction model for the Math post-course SAT score, and one background variable, Grade Point Average (GPA), was a significant covariate in predicting the Verbal SAT score. No interaction effects were significant in either model.

The model that best predicts the Verbal post-course SAT shows a significant effect of PSAT, Coaching and GPA. The model explains 71% of the variation in the Verbal post-course SAT. The effect of group membership is highly significant, and once Verbal PSAT score and GPA are controlled for, the effect of the coaching intervention on the Verbal test is estimated by the model to be 52.3 points ($p<.001$).

The significance of the GPA covariate was somewhat of a surprise, given that the means of the two groups were quite similar (Group 1 mean GPA = 87.2, Group 2 mean GPA = 87.8 ($t=.73$, n.s.)), and the correlation between GPA and Verbal SAT is moderate ($r=.34$). According to the model, a one point increase in GPA is associated with a 3.7 point increase in Verbal SAT scores even when the Verbal PSAT score and coaching have been accounted for ($p<.02$). This means that if two coached students have identical PSAT scores, the student getting better grades in school is expected to score higher on the Verbal section of the SAT. The GPA covariate explains 3.6% of the variation in post-course Verbal SAT scores over and above the variation explained by Verbal PSAT scores and coaching. The Verbal coaching effect without controlling for GPA is 49.3 points.

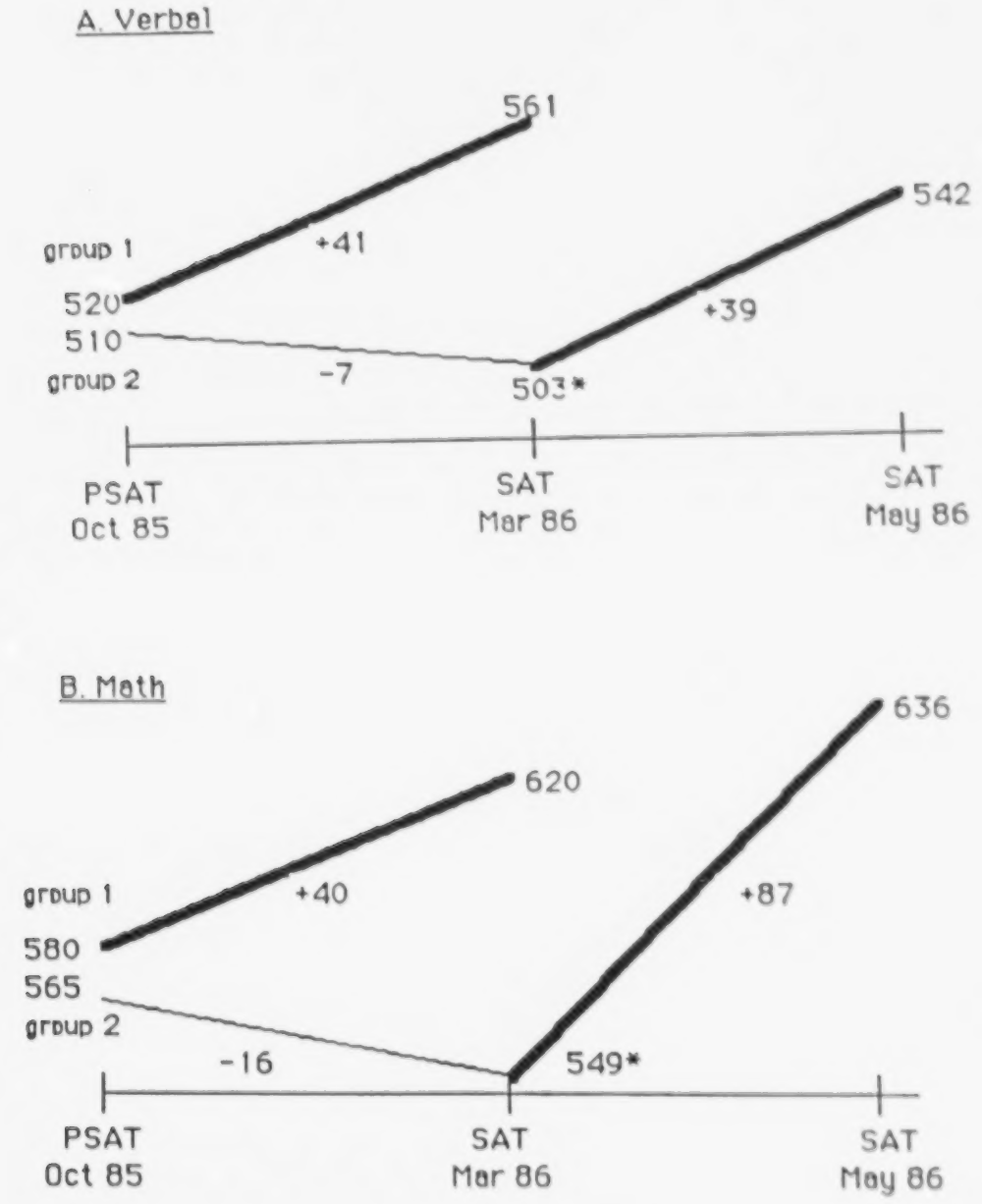
The model to predict the Math post-course SAT score shows a significant effect of PSAT and Coaching. No other covariates were significant in helping to explain variation in the Math SAT. The model explains 77% of the variation in the Math post-course SAT. The effect of coaching is highly significant, and once Math PSAT score is controlled for, the effect of the coaching intervention on the Math test is estimated by the model to be 58.4 points ($p<.001$).

In summary, the effect of the coaching course during the first period of the study was statistically highly significant. The model estimates for the effects of the coaching intervention on SAT scores are about 52 points Verbal and about 58 points Math for regular students in the study.

Control group students went on to be coached during the second period of the study and took the SAT in May. Coaching effects cannot be estimated with precision for these students because no further controls were used, but PSAT to SAT gains can be reported. This group of 34 students had mean gains of 32 Verbal and 71 Math points from the Fall 1985 PSAT to the May

Figure 2: Summary of Mean Changes in Verbal and Math Scores for Regular Students.

(Accompanying statistics included in Table 1 of the Appendix)



1986 SAT. According to Donlon (1984, p. 65), the expected growth from the PSAT to the SAT for students scoring at this level is about 10 points Verbal and 5 points Math. While it is clear that the students enrolled in the coaching course are not typical PSAT takers, it seems that the gains registered by the coached students far surpass the expectations that could be attributable to growth or other external factors.

Figure 2 illustrates the mean raw Verbal and Math gains of the regular students in the treatment and control groups. Accompanying stastistics are included in Table 1 of the Appendix. Note that the mean scores of the control group decreased somewhat from the PSAT to the pre-coaching SAT. If the two test administrations were equivalent, we would expect students' SAT scores to be somewhat higher than their PSAT scores due to growth and maturation.

Several factors may have influenced the decrease in SAT scores of the control group. The SAT may have been more difficult than the PSAT, or may have been scored based on somewhat different norms than the PSAT. It is also possible that some students inflated self-reported PSAT scores. I verified PSAT scores of 20% of the regular students and found only one discrepancy. However the coaching school personnel reported that some of their students typically tend to inflate self-reported PSAT scores. All of these factors are likely to influence students in the coached and control groups in the same way, and do not impact on the size of the coaching effect estimate.

It is possible that the lower SAT scores of the control group are due at least in part to decreased motivation among this group when taking the SAT test. If this were the case the coaching effect estimate would be inflated to some extent. However, the test was administered under closely supervised conditions and the results of the test were used to place students in the appropriate level of coaching class. The students were told this before the test and showed no indication that they were not fully applying themselves. There is no basis on which to assume that control students' scores were deflated due to low motivation.

## Effects of Coaching for Scholarship Students

A number of scholarship students did not complete the coaching course, therefore analyses are based on the 16 coached and 17 control students who completed the course. The substantial attrition rate (31%) and resultant small sample size limit the reliability of the scholarship study results, even though the treatment and control groups remained equivalent on 19 available background variables.

All scholarship students took the unreleased December 1985 version of the SAT at the end of the coaching course, and 14 of the 16 students in the treatment group took the May SAT the following weekend. For these 14 students the results of the two tests were nearly identical. The correlation between the two tests was very high (.92 Verbal, .89 Math), and the mean scores on the May SAT differed by only -6 Verbal points and +6 Math points from the test taken the previous weekend. Due to the seemingly high reliability of the December '85 SAT administered to all scholarship students, I used those scores as the posttest in the analyses.

Regression analyses predicting post-course SAT scores controlling for the PSAT showed no significant effect of coaching on Verbal scores, and a significant 57.0 point effect estimate of coaching on Math scores ($p<.001$). In other words, the coaching for scholarship students had no significant effect on Verbal SAT scores and about a 57 point effect on Math SAT scores. Due to the small sample size, a reliable analysis of background factors could not be carried out.

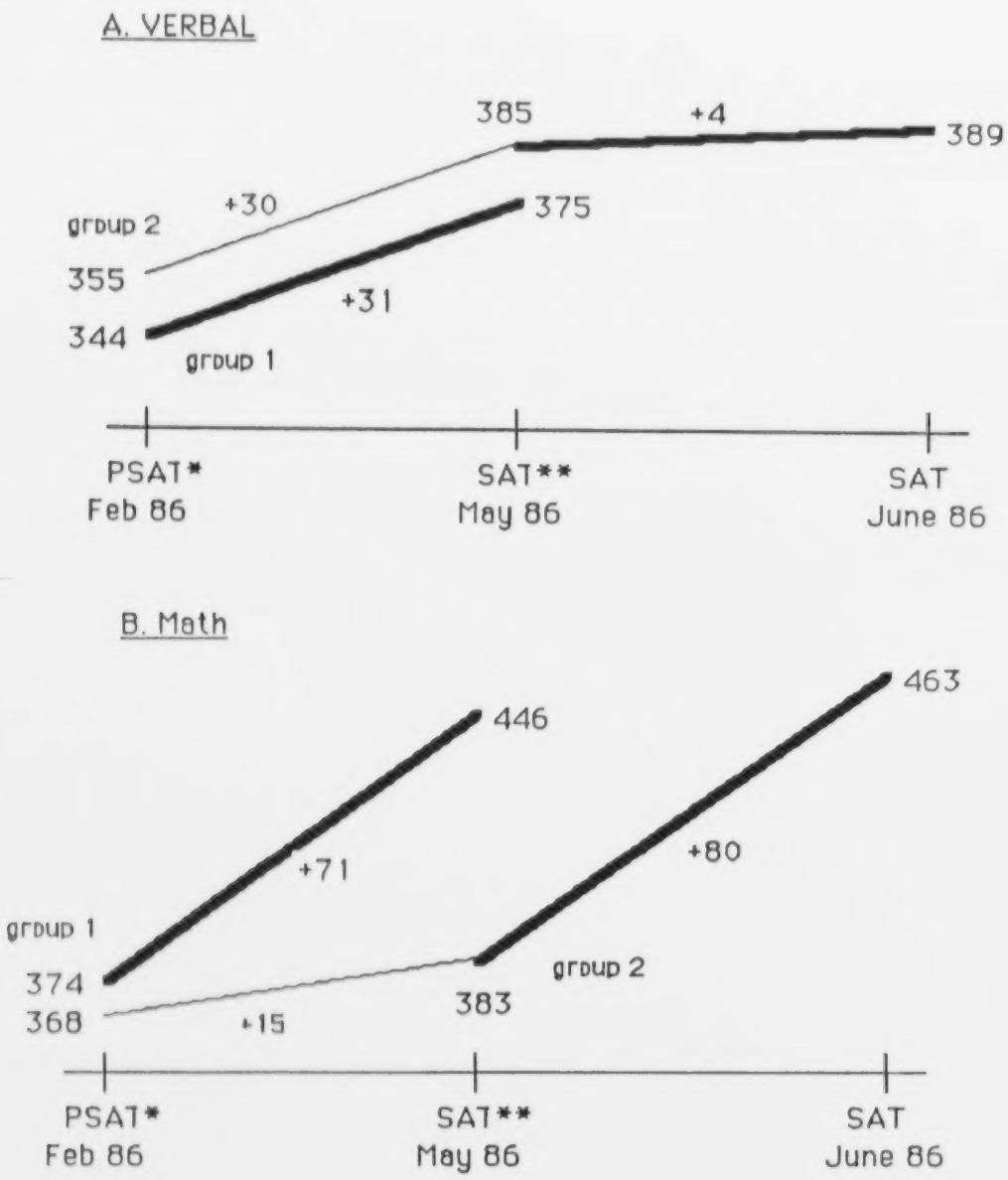
Scholarship students in the control group went on to be coached during the second period of the study and 12 of them took the SAT in June. All 17 also took the May SAT, supplied by the Educational Testing Service, at the conclusion of the coaching course the weekend before the June SAT. The June SAT scores for the five students who did not take the exam were estimated from the May SAT results. Coaching effects cannot be estimated with precision for these students because no further controls were used, but gains in SAT scores can be reported. The control students increased their Verbal scores by an average of 4 points after coaching, and increased their Math scores by an average of 80 points. These results are consistent with the first part of the scholarship study, which indicated no significant effect of coaching on Verbal scores, but a substantial effect on Math scores. The mean Verbal and Math gains for both scholarship student groups are illustrated in Figure 3. Accompanying statistics are included in Table 2 of the Appendix.

## Discussion

The significant gains in SAT scores achieved by students in this study indicates that coaching <u>can</u> be effective in inceasing students' scores. The question is no longer whether coaching can be effective, but what is it that makes some coaching effective. As DerSimonian and Laird (1983) have pointed out, commercial coaching courses, which have not previously been studied using experimental designs, have generally yielded higher average gains than ad-hoc or school-based preparation efforts. The results of this study support the conclusion that an intensive, well-planned course focused precisely on SAT-type

Figure 3: Summary of Mean Changes in Verbal and Math Scores for Regular Students.

(Accompanying statistics included in Table 2 of the Appendix)



questions can have a significant effect on increasing students' SAT scores.

The reason that this common-sense conclusion has not achieved a consensus in the literature is because of the major effort to combine the results of previous coaching studies, rather than trying to understand how the coaching programs that produced significant gains were different from programs that did not produce gains. Part of the problem has been that in general, the actual coaching interventions have not been well documented in the literature, and that the research designs of some studies did not preclude alternative explanations of gains in SAT scores. However the efforts by meta-analysts (DerSimonian and Laird, 1983; Messick and Jungeblut, 1981; Messick, 1982; Slack and Porter, 1980) to combine the results of effective and ineffective coaching interventions have clouded the central issue. There is evidence in the literature, for example Evans and Pike (1973), Kulik et. al. (1984) and Pike (1978), that focused instruction aimed at giving practice on SAT item-types, such as the type of instruction provided by some commercial coaching schools, can significantly enhance students' performance on the SAT. The results of this study certainly support such a conclusion.

How do the results of the present study compare with previous research on SAT coaching effectiveness? Messick (1982) analyzed the results of existing coaching studies according to contact time. He found a positive logarithmic relationship between coaching effect and contact time, with additional time spent on coaching for the SAT expected to yield more modest effects than an initial amount. The regular students in the present study had a total of approximately 16 hours of class preparation. According to Messick's research, these students would be expected to increase their Verbal scores due to coaching by 11 points and their Math scores by 17 points. If homework and extra help are included in contact time for a total of 32 hours, Messick's research predicts a coaching effect of 16 Verbal points and 25 Math points. Clearly the 52 point Verbal and 58 point Math coaching effects estimated in the present study for regular students are much larger than Messick would predict. According to Messick's data, such gains might be expected with contact time of about 300 hours for Math, and are not predicted at all for Verbal. This comparison is particularly important because the College Board's official statement about the effect of coaching is based on the research that Messick used in his analyses.

What are some of the reasons that the coaching program was differentially effective for regular and scholarship students? One reason is probably that the coaching school staff had no prior experience teaching minority students. The instructors had to change their methods to meet the needs of the scholarship students, and the results were not always successful. There was

some indication that the Verbal preparation given by the coaching school may have confused some of the scholarship students. A few of the scholarship students who responded to the post-course questionnaire said that they tried to use some of the approaches to the Verbal test questions taught by the coaching school, but that they did not find them useful. One scholarship student who withdrew from the study wrote me a letter in which she said that the reason that she decided to stop attending the coaching course was that she noticed that her Verbal scores were going down, and that the approach taught by the coaching school did not work for her.

Some of the scholarship students felt that they were not adequately prepared to take the real SAT at the conclusion of the course. This was echoed by the instructors, who all agreed that they would have needed more time and a more intense course schedule in order to increase the Verbal scores of the scholarship students. The instructors pointed out that many of these students come from families where English is a second language and their vocabulary was therefore limited. They also said that some of the Black students in the study lacked the vocabulary base necessary to deal with SAT-type Verbal questions. The instructors emphasized repeatedly that the scholarship students who participated in the study were all bright, motivated individuals who could all do college-level work if they were given adequate preparation. The instructors expressed anger and frustration at seeing that these students were not being given the same quality of instruction and preparation in their high schools as their regular students received. They all agreed that these scholarship students were being denied opportunities by the lack of attention and education that they received in school, and they felt frustrated because they could not help these students more.

The instructors worked with both populations, and they could see and experience first-hand some of the differences and inequalities between the two groups. The instructors learned that the scholarship students in the study, despite their low average SAT scores, were some of the best students in their Manhattan and Brooklyn high schools. They realized that these students were simply not being taught the basic verbal and math skills necessary to do well on the SAT. Their overwhelming response was nothing short of an indictment of the poor quality of education in low-income, minority urban schools. The instructors believed that the scholarship students were fully capable of being successful in college, but the poor quality of education they received in their schools prevented them from achieving their potential.

Such a conclusion is fully backed by recent research into the plight of poor and minority students. One example is the report of the Board of Inquiry commissioned by the National Coalition of Advocates for Students, which is entitled "Barriers to Excellence: Our Children at Risk" (1986). The report

concludes that new efforts are needed to prevent poor urban and minority students from suffering increasing inequalities in educational opportunity.

Such a conclusion sheds new light on efforts to coach minority students for tests like the SAT. While such efforts may help individual minority students gain access to college, they do not address the underlying issue of the poor quality of education that these students are receiving. In the present study the average combined PSAT scores of the regular student sample was more than 350 points higher than for the scholarship student sample. This difference is one indicator of the discrepancy in the quality of schooling experienced by the two groups. It will take much more than a coaching course to diminish this discrepancy.

This research is a pilot effort to assess the effect of a commercial coaching school on two groups of students at one site. The study is limited by a small sample size and attrition from the scholarship student sample, as well as the non-randomized assignment of the regular students to treatment and control groups. The small sample size and the problems of implementation limit the generalizability and reliability of the findings. The results are suggestive, but larger, well-controlled studies are necessary to substantiate the coaching effects that were found.

The focus of future research in the area should be on what makes some coaching effective and on what students learn when they are coached effectively. The results of the present study, along with results of some previous studies of coaching, indicate that coaching can be effective in increasing SAT scores. What needs further investigation is what type of coaching is particularly effective. There is some indication that coaching is more effective when it is focused specifically on practice and understanding of SAT-type items. There is also some indication that some commercial coaching programs, which can bring greater resources and intensity to special preparation, may be more effective in increasing students' SAT scores. These tendencies should be researched more thoroughly in order to be fully substantiated.

Additionally, future research needs to examine if effective coaching influences the quality of student work in college. If coaching that results in increased SAT scores does not result in comparable increases in scholastic aptitude, then the predictive validity of the SAT for coached students in called to question. Responses to the questionnaire administered at the conclusion of the coaching course indicate that students perceived that the coaching increased their SAT scores, but did not affect the quality of their school work. The result is preliminary, but suggests that more research on the predictive validity of the SAT for coached students would be appropriate.

## BIBLIOGRAPHY


Alderman, D.L., and Powers, D.E., "The Effects of Special Preparation on SAT-Verbal Scores", ETS Research Report RR-79-1, Princeton NJ:ETS, 1979.

DerSimonian, R., and Laird, N.M., "Evaluating the Effect of Coaching on SAT Scores: A Meta-Analysis", Harvard Educational Review, Vol. 53, No. 1, Feb. 1983, pp. 1-15.

Donlon, T.F., editor, The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests, NY:CEEB, 1984.

Evans, F.R., and Pike, L.W., "The Effects of Instruction for Three Mathematics Item Formats", Journal of Educational Measurement, No. 10, 1973, pp. 257-271.

Federal Trade Commission, Boston Regional Office, Staff Memorandum of the Boston Regional Office of the FTC: The Effects of Coaching on Standardized Admission Examinations, Boston MA:FTC, September 1978.

Federal Trade Commission, Bureau of Consumer Protection, Effects of Coaching on Standardized Admission Examinations. Revised Statistical Analyses of Data Gathered by the Boston Regional Office of the FTC, Washington DC:FTC, March 1979.

Kulik, J.A., Kulik, C.C., and Bangert, R.L., "Effects of Practice on Aptitude and Achievement Test Scores", American Educational Research Journal, Vol. 21, No. 2, Summer 1984, pp. 435-447.

Messick, S., The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research from the Fifties to the FTC, Princeton NJ:ETS, 1980.

Messick, S., "Issues of Effectiveness and Equity in the Coaching Controversy: Implications for Educational and Testing Practice", Educational Psychologist, Vol. 17, No. 2, 1982, pp. 67-91.

Messick, S., and Jungeblut, A., "Time and Method in Coaching for the SAT", Psychological Bulletin, Vol. 89, No. 2, 1981, pp. 191-216.

National Coalition of Advocates for Students, Barriers to Excellence: Our Children at Risk, Boston MA:NCAS, 1986.

Pike, L.W., "Short-term Instruction, Testwiseness, and the Scholastic Aptitude Test: A Literature Review with Research Recommendations", ETS Research Bulletin RB-78-2, Princeton NJ:ETS, 1978.

Powers, D.E., "Estimating the Effects of Various Methods of Preparing for the SAT", College Board Report No. 82-2, NY:CEEB, 1982.

Slack, W., and Porter, D., "The Scholastic Aptitude Test: A Critical Appraisal", Harvard Educational Review, Vol. 50, No. 2, May 1980, pp. 154-175.

Zuman, J. P., "The Effectiveness of Special Preparation for the SAT: An Evaluation of a Commercial Coaching School", Doctoral Dissertation, Cambridge MA:Harvard Graduate School of Education, June 1987.

## APPENDIX

For a complete description of the research methodology, discussion of analyses, and presentation of further results of this study, please refer to Zuman (1987).

Table 1: Means and Standard Deviations of Test Scores for Regular Students

| | GROUP 1 n=21 | | GROUP 2 n=34 | |
|---|---|---|---|---|
| | Mean | S. D. | Mean | S.D. |
| **VERBAL** | | | | |
| PSAT (Oct 85) | 520 | 85.8 | 510 | 83.6 |
| SAT (Mar 86) | 561 | 78.5 | 503* | 102.4 |
| SAT (May 86) | - | - | 542 | 80.2 |
| **MATH** | | | | |
| PSAT (Oct 85) | 580 | 131.3 | 565 | 90.5 |
| SAT (Mar 86) | 620 | 102.5 | 549* | 96.8 |
| SAT (May 86) | - | - | 636 | 82.1 |

* Group 2 March SAT scores are adjusted from special SAT test.

Table 2: Means and Standard Deviations of Test Scores for Scholarship Students

| | GROUP 1 n=16 | | GROUP 2 n=17 | |
|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. |
| **VERBAL** | | | | |
| PSAT (Feb 86) | 344 | 59.3 | 355 | 108.2 |
| SAT (May 86) | 375 | 56.7 | 385 | 73.3 |
| SAT (June 86) | - | - | 389 | 86.1 |
| **MATH** | | | | |
| PSAT (Feb 86)* | 374 | 74.0 | 368 | 77.5 |
| SAT (May 86)** | 446 | 104.7 | 383 | 82.9 |
| SAT (June 86) | - | - | 463 | 107.2 |

* Released Nov. 85 version of PSAT
** Dec. 85 version of SAT used by permission of ETS